

# **L'extraction de relations temporelles cliniques**

avec la régularisation de la logique souple probabiliste et l'inférence  
globale

Hao ZHANG

# 1. Contexte

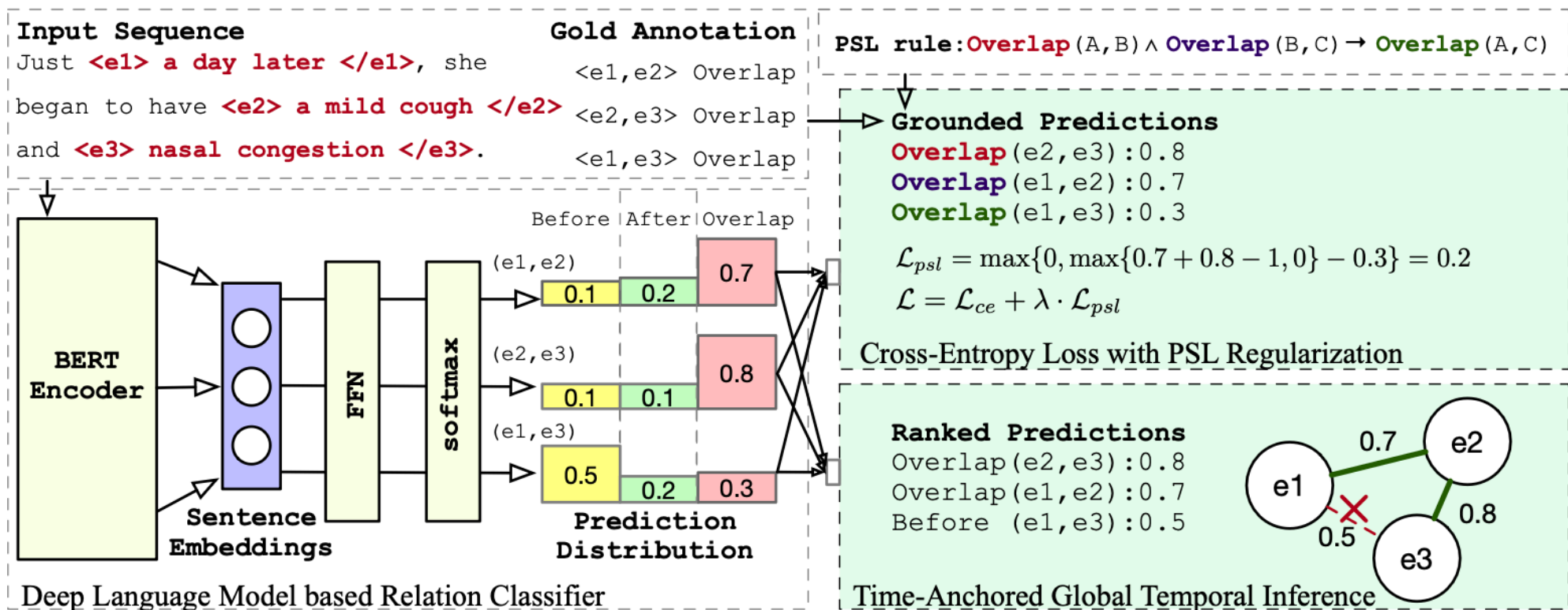
- Les rapports de cas cliniques jouent un rôle essentiel dans le partage d'expériences cliniques particulières et de phénotypes de maladies atypiques.
- L'extraction des relations temporelles entre les événements cliniques est cruciale pour les systèmes d'extraction de rapports de cas et de réponse aux questions médicales.
- Les méthodes existantes se trouvent en difficulté pour capturer les dépendances relationnelles globales et requièrent une ingénierie des fonctionnalités étendue.

## 2. Problématique

- La communauté médicale a un besoin persistant d'extraire avec précision les relations temporelles des récits cliniques.
- Les approches actuelles ne permettent pas de modéliser les dépendances temporelles complexes au sein des documents.
- Une méthode permettant de capturer et d'inférer efficacement les relations temporelles au niveau des documents est essentielle pour faire progresser la recherche et les applications dans le domaine des sciences médicales.

# 3. Méthode proposée : CTRL-PG

- La méthode proposée, CTRL-PG, vise à répondre aux limites des approches existantes en s'appuyant sur les règles de la logique souple probabiliste (PSL).
- En incorporant la régularisation PSL et l'inférence globale, CTRL-PG offre un nouveau cadre pour l'extraction précise des relations temporelles dans les textes médicaux.
- Motivation : améliorer l'efficacité et la précision de l'extraction de relations temporelles, permettant des applications plus avancées dans les soins de santé et la recherche médicale.



# 3. Méthode proposée : CTRL-PG

- Composants : Classificateur de relations temporelles, fonction de perte à entropie croisée avec régularisation PSL, module d'inférence temporelle globale à ancrage temporel.

## 4. Classificateur de relations temporelles

- Le classificateur de relations temporelles de CTRL-PG est un composant crucial qui utilise un encodeur de langage profond, tel que BERT, pour comprendre le contexte d'un texte clinique.
- Il extrait des caractéristiques, les traite par le biais d'un réseau en amont (Feed-Forward Network - FFN) et prédit les relations temporelles entre les événements cliniques.
- En attribuant des étiquettes de relations temporelles appropriées et en intégrant la régularisation PSL, le classificateur renforce la capacité du modèle à identifier et à classer avec précision les relations temporelles, améliorant ainsi l'extraction globale d'informations temporelles à partir de documents cliniques.

## 5. Fonction de perte à entropie croisée avec régularisation PSL

- La régularisation dans CTRL-PG combine le calcul de perte traditionnel avec les règles PSL pour améliorer les performances du modèle dans l'extraction de relations temporelles.
- En pénalisant les instances d'apprentissage qui violent les règles de dépendance temporelle prédéfinies codées sous forme de règles PSL, le terme de régularisation garantit que le modèle apprend à adhérer à ces règles au cours de l'apprentissage.
- Cette approche permet une modélisation plus souple et plus efficace des dépendances temporelles, ce qui améliore la précision et la robustesse de la classification des relations temporelles entre les événements cliniques.

## 6. Module d'inférence temporelle globale ancrée dans le temps

- Le module d'inférence temporelle globale ancrée dans le temps de CTRL-PG s'appuie sur l'algorithme Timegraph pour résoudre les conflits dans les prédictions de relations temporelles.
- En construisant un graphe temporel sans conflit pour chaque document grâce à un processus de vérification et d'ajout, le module détermine efficacement les relations temporelles entre les entités nommées sur la base des prédictions et de leurs probabilités associées.
- Ce mécanisme d'inférence globale améliore la capacité du modèle à faire des prédictions de relations temporelles précises et cohérentes au niveau du document, améliorant ainsi les performances globales de l'extraction de relations temporelles dans l'analyse de textes cliniques.



# 7. Principales contributions

- Introduction de la régularisation de la logique souple probabiliste (PSL) pour renforcer les règles de dépendance temporelle pendant l'apprentissage du modèle.
- Implémentation d'un module d'inférence temporelle globale utilisant l'algorithme Timegraph pour des prédictions de relations temporelles précises.
- Proposition de l'architecture du modèle CTRL-PG qui combine un classificateur de relations temporelles contenant un encodeur de langage profond, une régularisation PSL et une inférence globale.
- Améliorer l'apprentissage du modèle avec une fonction de perte d'entropie croisée et une régularisation PSL pour capturer les dépendances temporelles complexes.
- Mise à disposition de codes pour soutenir la recherche et le développement ultérieurs.

# 8. Résultats expérimentaux

- Expériences approfondies sur les ensembles de données I2B2-2012 et TB-Dense.
- Les résultats expérimentaux démontrent que le modèle CTRL-PG est nettement plus performant que les méthodes de référence pour l'extraction de relations temporelles cliniques sur les ensembles de données de référence.
- En incorporant la régularisation de la logique douce probabiliste (PSL) et un module d'inférence temporelle globale, le modèle CTRL-PG montre des améliorations en termes de précision et d'efficacité.
- Malgré des défis tels que la gestion du bruit et l'impact des règles PSL, le modèle est prometteur dans la capture des dépendances temporelles complexes et l'amélioration des valeurs de rappel.

# 9. Conclusion

- Le modèle CTRL-PG, qui utilise la régularisation de la logique douce probabiliste et l'inférence temporelle globale, surpasse les méthodes de base dans l'extraction de relations temporelles cliniques.
- Malgré des défis tels que le bruit et l'impact des règles PSL, le modèle est prometteur dans la capture des dépendances temporelles complexes et l'amélioration des valeurs de rappel.
- Dans l'ensemble, CTRL-PG présente une avancée significative dans l'extraction précise des relations temporelles à partir de textes cliniques, améliorant ainsi l'efficacité de l'analyse narrative des soins de santé.

# 10. Orientations futures

- En termes d'orientations futures, la recherche pourrait se concentrer sur l'amélioration du modèle CTRL-PG en relevant les défis liés à la gestion du bruit et à l'impact des règles de la logique souple probabiliste (PSL).
- En plus, il serait bénéfique d'explorer des méthodes pour améliorer la performance du modèle en capturant des dépendances temporelles complexes et en améliorant les valeurs de rappel.
- En outre, l'étude de l'évolutivité et de l'applicabilité du modèle CTRL-PG à des ensembles de données plus importants et à diverses sources de textes cliniques pourrait fournir des informations précieuses.
- Les orientations futures pourraient consister à affiner l'architecture du modèle, à optimiser les techniques de formation et à explorer de nouvelles approches pour faire progresser le domaine de l'extraction des relations temporelles cliniques, etc.

**Merci**